
CD-1.1 Reliable Multicasting: Status and Plans

Ray Cordova, SAIC

Roger Bowman, SAIC

Deb Agarwal, LBNL

Gus Gustafson, Consultant

October 22, 2002

Topics

- Motivation
- Background
- Requirements
- Design
- CD-1.1 Formats and Protocols
- Plans
- Summary

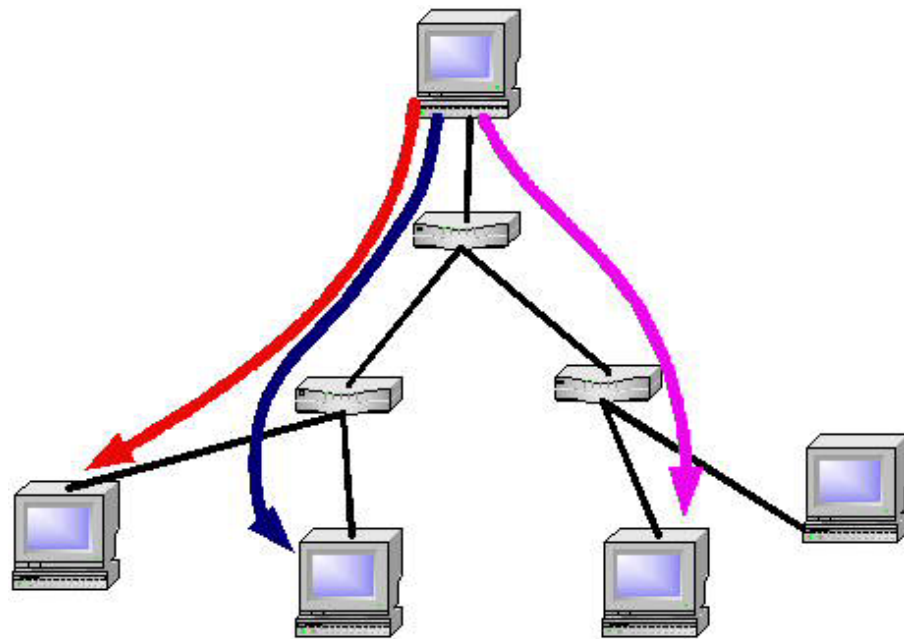
Motivation for CD-1.1 Reliable Multicast

- Enhance reliability of IMS data transmission to signatories
- Minimize reliability and data distribution burden (= Cost) for the IDC
- Provide direct transmission of IMS station data to NDC hosting station and IDC
- Minimize bandwidth required for transmission of data from a single data producer to multiple destinations

What is Multicast Communication?

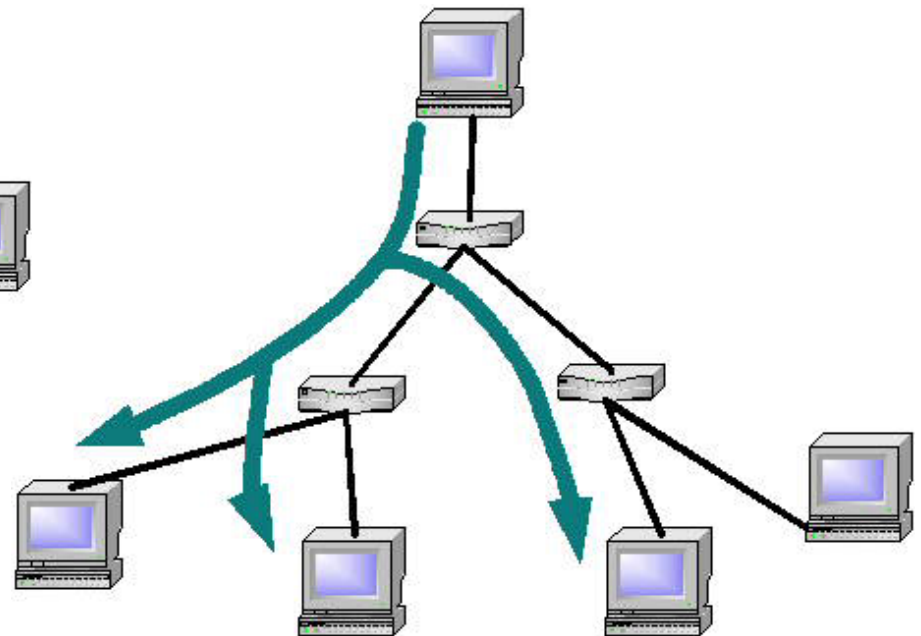
- Group communication mechanism
 - Provides one-to-many and many-to-many communication
- Efficient dissemination of messages
 - Network-based duplication (when needed)
 - Multicast retransmissions
 - Bandwidth savings
 - Parallel delivery at multiple locations

IP Multicast Communication



Unicast

Multicast



Internet Protocol (IP) Multicast

- Efficient group communication mechanism
 - Send a message to a group of receivers
 - Best-effort delivery to the group members
- Implemented in the routers and hosts
 - Class D addresses used for multicast (224.x.x.x - 239.x.x.x)
 - Network components manage routing and duplicate the message as needed
 - Co-exists with TCP and UDP communication mechanisms

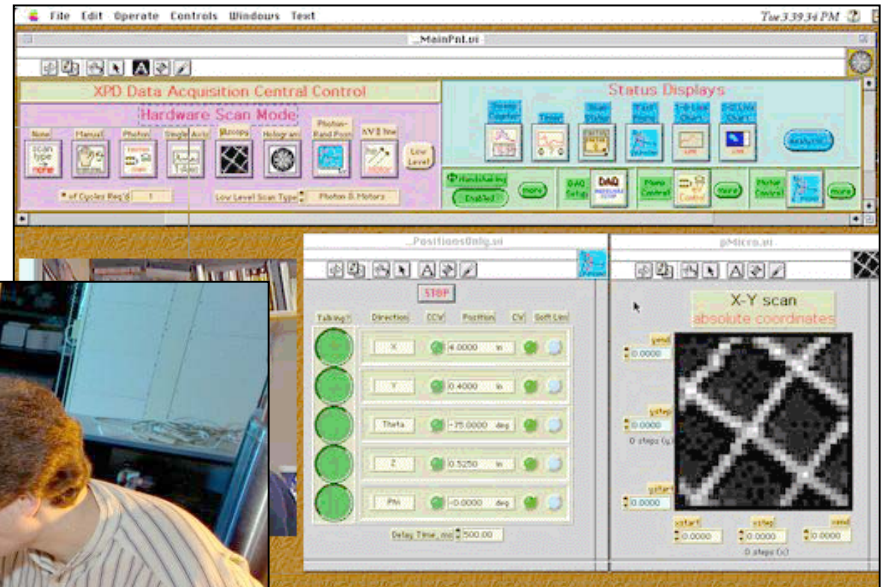
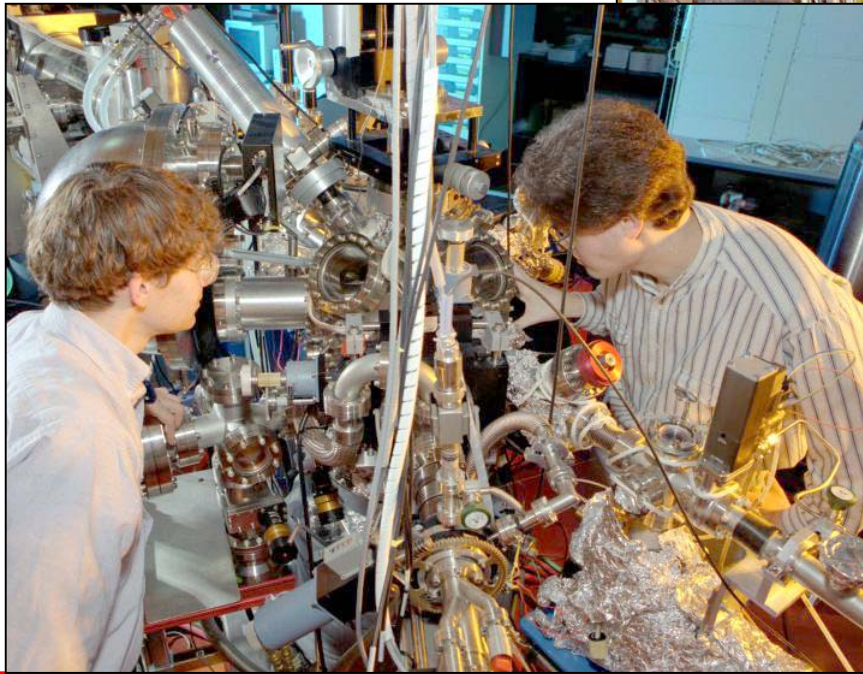
Example IP Multicast Use (Access Grid)



What is Reliable Multicast?

- Properties similar to TCP
- Application-level program
- Uses IP Multicast as the underlying communication mechanism
- Reliable and ordered delivery of messages within a group
- Tracks group membership
- IETF Reliable Multicast Transport Working Group is defining standardized building blocks

Example Reliable Multicast Use (Remote Instrument Access)



Reliable Multicast Goal: TCP of Multicast

- What is unicast TCP?

Internet Layer	Unicast Protocols	
Transport Layer 4	(Stream): TCP	(Datagram): UDP
Network Layer 3	IP	

- What does unicast TCP do?
 - Reliable unicast transport of a data stream: **most recognized, but not only function**
 - Congestion control (helps protect network from overload)
 - Flow control (match rate of sender and receiver)
 - Synchronize start of data stream for sender and receiver
- Note: these problems are much harder for multicast

Terminology

Term	Definition	Service
AEC	Application-level Error Correction	Tracking of data that have been sent and received, and retransmission of missing data as necessary
Multicast	Group delivery mechanism	Best-effort delivery
Reliable multicast	Group delivery with packet-level AEC	Reliable delivery of data packets absent major failure modes (better best effort)
CD-1.1 reliable multicast	Group delivery with packet-level AEC and individual delivery with frame-level AEC	Reliable delivery of Data Frames, even with major failure modes
Reliability host	Application program providing re-transmission of data packets by multicast or Data Frames by point-to-point mechanisms	Delivery of data packets or Data Frames that are missing at one or more data consumer.

CD-1.1 Reliable Multicast

- Assumptions
 - Reliability of CD-1.1 multicasting should be equivalent to that of CD-1.1 unicasting
 - Bandwidth of tail circuit to stations will not be increased to support multicast (in catch-up mode)

Requirements Overview

- System-level requirements are a minor extension of Continuous Data Subsystem (CDS) CD-1.1 unicast requirements
 - As documented in CDS software design description [IDC7.4.1] (July 2001)
- New requirements are defined for two subsystems
 - Multicast data transmission subsystem, including packet-level application error correction
 - Frame-level application error correction subsystem by Point-to-point data transmission

System-Level Requirements

- All system-level CD-1.1 unicast requirements apply
 - 99.99% reliability requirement has been retained
 - Note that this was tightened from 99.9% in [IDC7.4.1]
 - Application level reliability provided by a combination of multicast and point-to-point transmission mechanisms
 - The connection requirement has been modified to generalize initiation of a connection after an outage
- New system-level requirement
 - “The CDS shall provide multicasting capability to deliver CD-1.1 frames using underlying IP router multicast capability.”

Design Requirements for the Multicast Capabilities

- Multiple data providers shall be capable of providing multicast data over a transport network.
- Multicasting shall support all IMS stations in the Treaty (up to 321 stations) in multicasting data over a transport network.
- Transmission rates for multicast data shall be configurable to mitigate network congestion.
- A multicast data provider shall support up to 20 data consumers in a single multicast group.
- A multicast data provider shall not be limited in its sending by the absence or presence of any specific multicast data consumer.

Multicast Requirements (2)

- Multicasting shall support increases and decreases in the size of the multicast group without the need to restart the sending activity of the data provider.
- The size of multicast data packets shall be configurable to support the smallest MTU (maximum transmission unit) used by the transport network.
- A method for identifying, requesting, and resending missing multicast data packets of an active multicast session shall be provided, limited by configurable data buffer size.
- A multicast data provider shall transmit frames in the order in which they were written to its CD-1.1 frame set, with the exception of retransmissions.

Multicast Requirements (3)

- A multicast data provider shall begin delivery at the current time or a configurable lookback time less than 10 minutes prior to the current time to prevent gaps in the stream caused by short data provider outages.
- A multicast data consumer shall be capable of entering or leaving a multicast group without negatively impacting other group members.
- Multicasting shall require less bandwidth on the data-provider communication link to service four multicast data consumers than required by CD-1.1 point-to-point communications from a data provider directly to two point-to-point data consumers.
- A multicast group shall correspond to one and only one frame set.

Point-to-Point Recovery Requirements

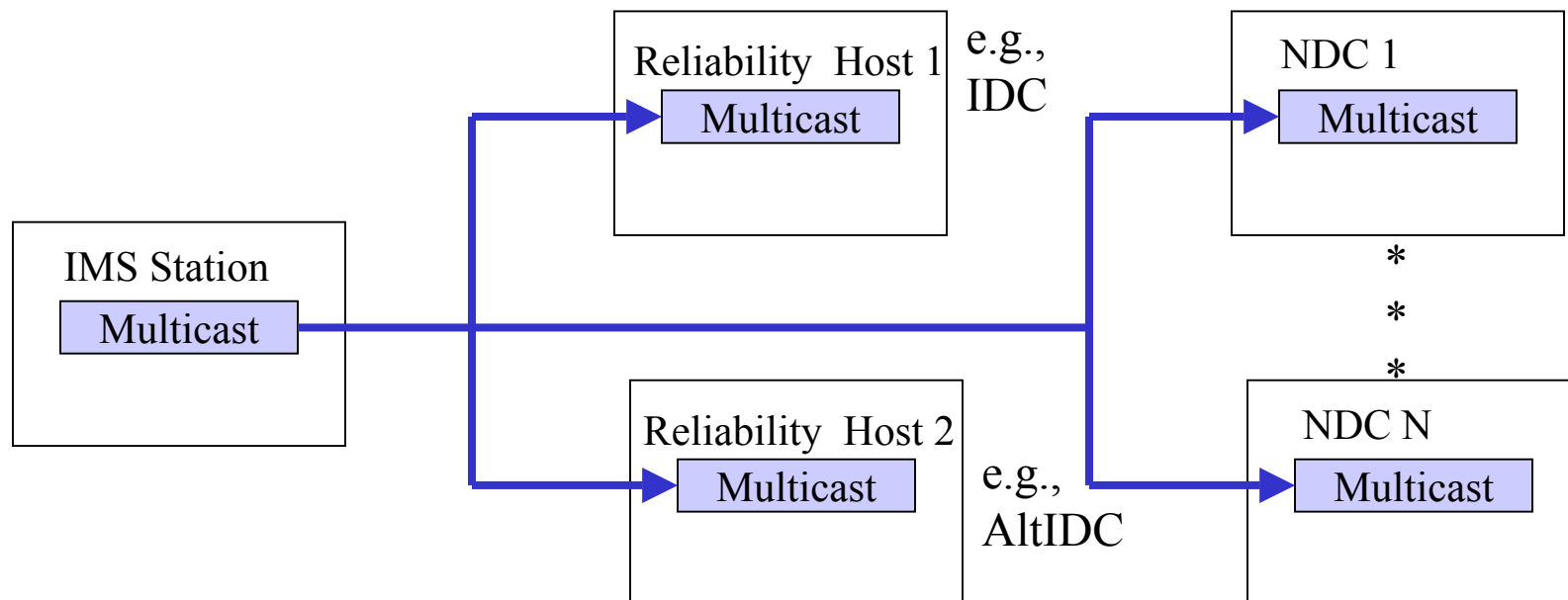
- The CDS shall use sequence numbers to detect frames not received by a data consumer via multicast.
- The CDS shall attempt to reliably deliver to a data consumer frames not received via multicast.

Design Constraints

- Minimize perturbations to CD-1.1 Formats and Protocols [IDC 3.4.3 Rev. 0.2]
- Compatibility with CDS CD-1.1 unicast
 - A data provider must be able to service both multicast and point-to-point data consumers simultaneously
- CDS CD-1.1 software components to operate in either multicast or unicast mode, according to parameter settings

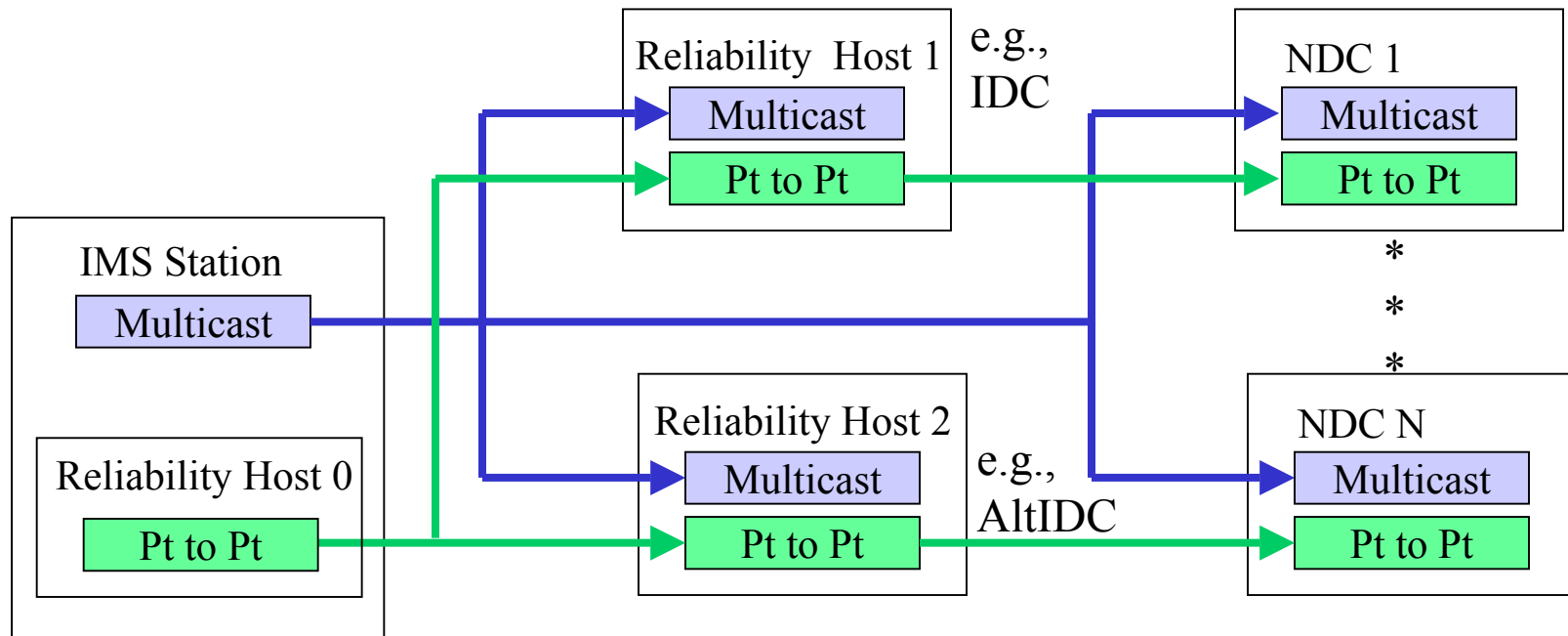
CD-1.1 Reliable Multicast Participants

An Example - Normal Operation

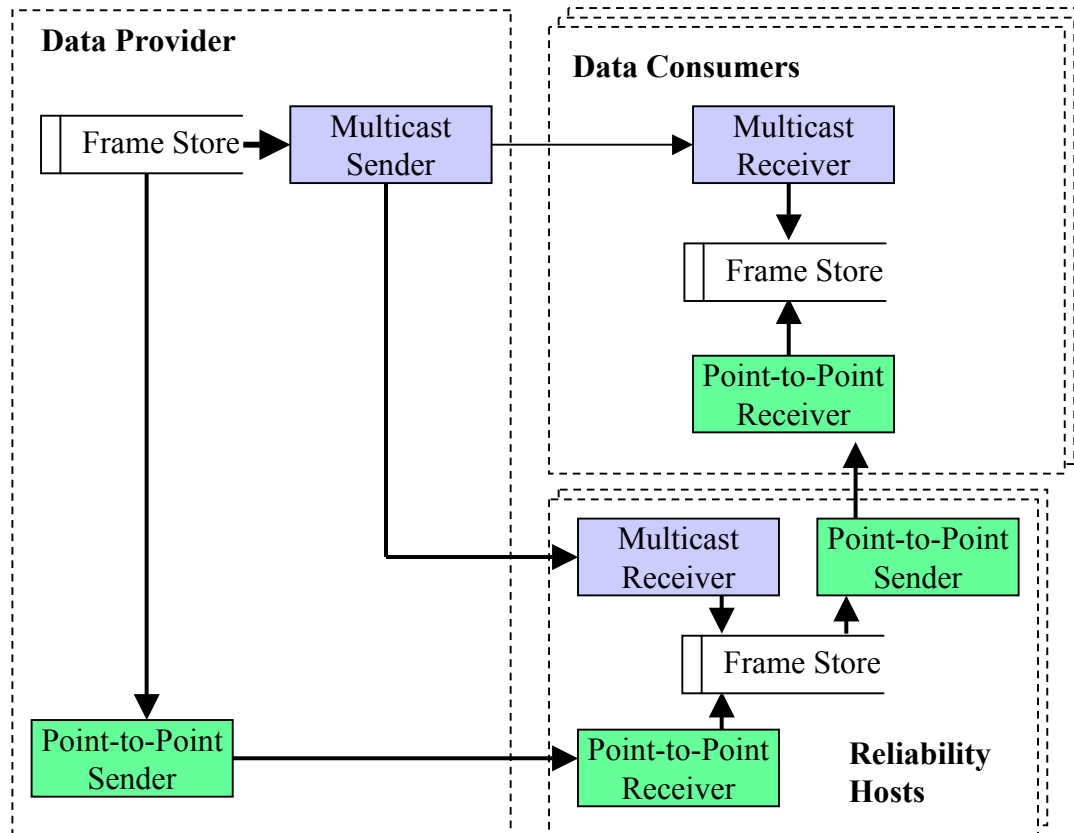


CD-1.1 Reliable Multicast Participants

An Example - Catch-Up Operation



Conceptual Design



Reliability Hosts

- Policies
 - IDC may provide reliability host for catch-up data transmission
 - NDCs may select location of reliability host (from short list)
 - All NDCs are entitled to the same level of reliability

Design Approach

- System level
 - Use custom reliable multicast solution (based on CD1.1)
 - Use point-to-point mechanism to provide application level reliability
 - Separate multicast (real-time) and point-to-point (catch-up) transmission into separate subsystems
- Multicast subsystem
 - Multicast transmission initiated by pull from data consumer
 - Use CD-1.1 procedure (in reverse) to establish connection
 - Multicast transmission begins at “data provider time” (current time - small look back) - no attempt to catch up
 - Data provider provides packet-level reliability host

CD-1.1 Formats and Protocols

- Extension to connection options
 - Data consumer requests connection for multicast transmission and for point-to-point to provide missing frames
 - Not changed for stations not multicasting
- Minor changes required to [IDC 3.4.3 Rev. 0.2]
 - Will use some fields that had been reserved for multicasting
 - Will define a new type of Option Request Frame
 - Will change miscellaneous text descriptions to be valid for both unicast and multicast

Response to Failure Modes

Failure mode	Design response
Packet loss (normal mode)	<ol style="list-style-type: none">1. Consumer sends PNack(s).2. Provider re-multicasts missing packet(s).
Network congestion (normal mode)	<ol style="list-style-type: none">1. Provider is limited to a maximum UDP transmission rate, which will be set to account for predicted TCP traffic.
Station or tail circuit outage	<ol style="list-style-type: none">1. All multicast sessions disconnect.2. Consumers initiate multicast reconnections with Provider.3. Provider begins to multicast at current time minus small lookback.4. If outage duration exceeds lookback, Consumer identifies missing Data Frames and initiates point-to-point reconnection with Reliability Host.5. Consumer pulls missing Data Frames from Reliability Host.

Response to Failure Modes (2)

Failure mode	Design response
Reliability Host outage	<p>No consequence to Consumers if multicast and packet reliability service from Provider operates without error. If a Data Frame has been determined to be missing, then:</p> <ol style="list-style-type: none">1. Consumers keep trying to connect with Reliability Host until successful.2. Reliability Host connects to Provider and pulls data using point-to-point mechanism.3. Consumers pull missing Data Frames from Reliability Host.
Regional network node outage	<ol style="list-style-type: none">1. Multicast sessions with affected Consumers disconnect.2. Affected Consumers initiate multicast reconnection with Provider.3. Affected Consumers initiate point-to-point reconnection with their Reliability Hosts.4. Consumers pull missing Data Frames from Reliability Hosts.

Response to Failure Modes (3)

Failure mode	Design response
Data Consumer or tail circuit outage	<ol style="list-style-type: none">1. Multicast session with Provider disconnects.2. Consumer initiates multicast reconnection with Provider.3. Consumer initiates point-to-point reconnection with its Reliability Host.4. Consumer pulls missing Data Frames from Reliability Host.

Design Issues

- Pull model for multicast transmission connection
 - (Pull model for point-to-point is not an issue)
- Negotiation/publication of Formats & Protocols revisions

Plans

- Finalize design
 - Examine failure modes and design response
 - Conduct design review
 - Update CD-1.1 Formats and Protocols document for multicast design
- Implement new multicast software components
- Adapt existing components of Continuous Data Subsystem CD-1.1, as necessary

Plans (2)

- Plan and conduct wide-area multicast testing
 - Phase 1:
 - Data provider at SAIC Testbed (San Diego)
 - Data consumers at CMR and SAIC Testbed
 - First quarter of 2003
 - Phase 2:
 - Data provider at I57US
 - Data consumers at CMR and SAIC Testbed or US NDC
 - Second quarter of 2003
 - Possible Phase 3:
 - Partitioned Subnetwork Configuration -
 - IMS multicast station to host NDC and IDC

Summary

- Requirement for reliable data delivery and bandwidth constraint is difficult but surmountable challenge
- Policy to avoid IDC single point failure requires flexibility in providing reliable service, and imposes higher floor on minimum bandwidth necessary (e.g., allow outage catch-up to IDC while continuing multicast of realtime data)
- Two-component design:
 - Multicast provides real-time transmission
 - Point-to-point provides reliable catch-up
- Wide-area proof-of-concept tests planned for 2003